Monte Carlo generator tuning for cosmic-ray induced air shower simulations

Workshop on the tuning of hadronic interaction models

Kevin Kröninger, **Michael Windau**

TU Dortmund, AG Kröninger

24.01.24

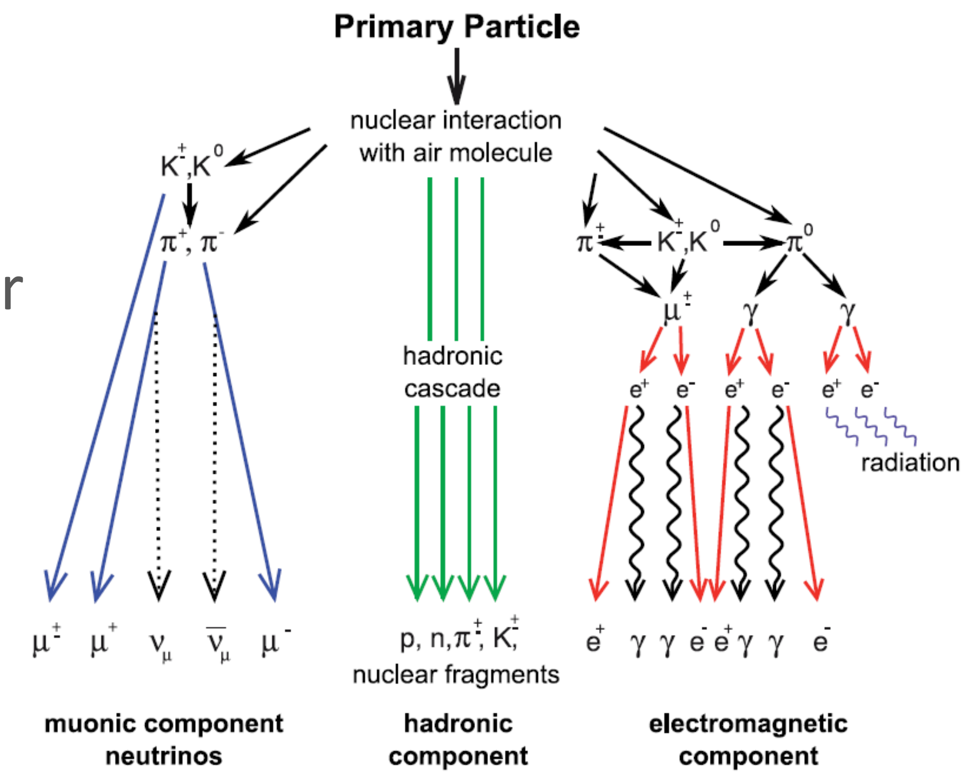
The need for MC generator tuning

Hadronization of process cannot be calculated from first principle
→ Need for hadronic models (MC generators)

Hadronization process has large impact on air shower features

Tuning essential to achieve high-quality simulations

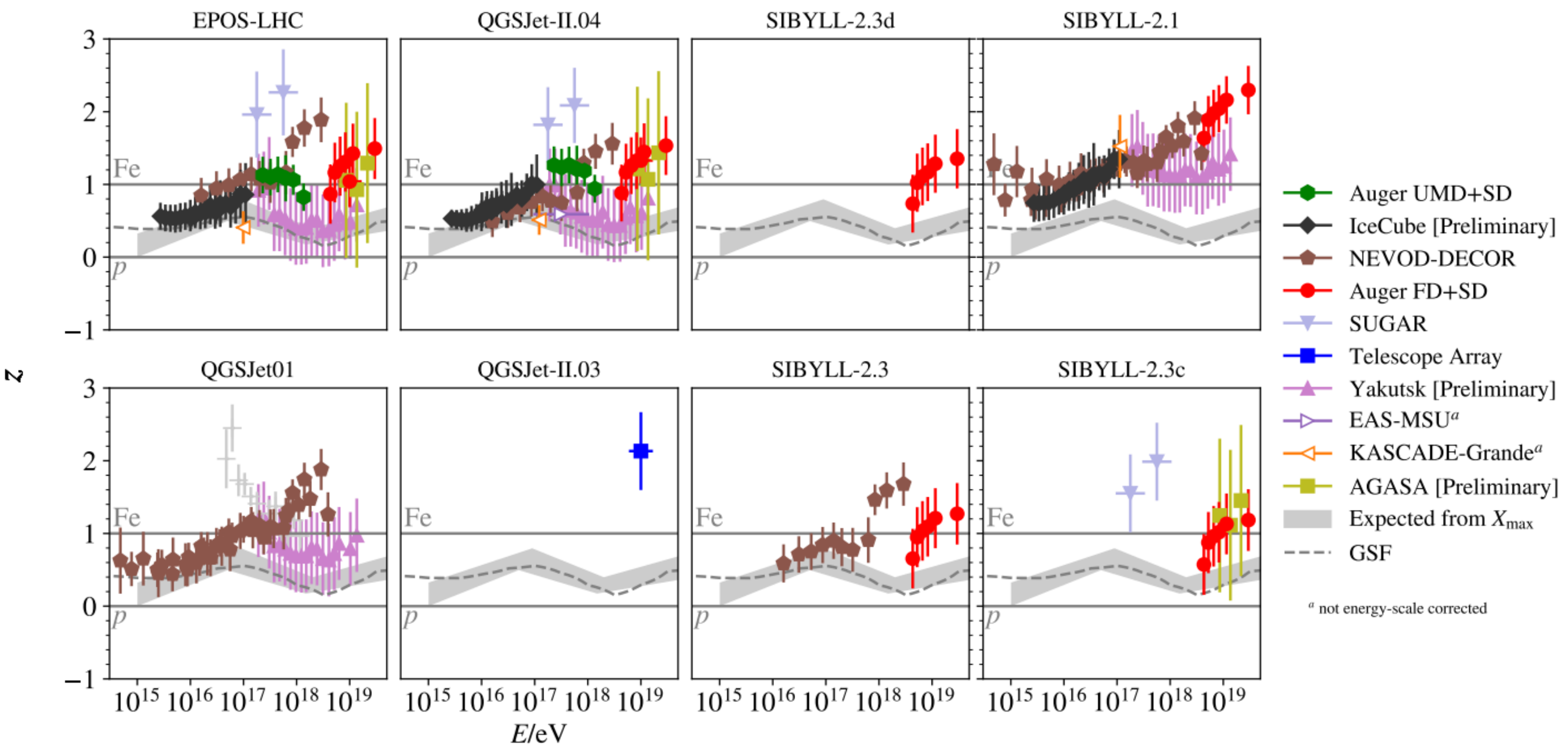
→ Muon Puzzle



Brief Reminder: Muon Puzzle

High energies: Muon excess in observations compared to simulations

$$z = \frac{\ln(N_{\mu}^{\text{det}}) - \ln(N_{\mu,p-\text{sim}}^{\text{det}})}{\ln(N_{\mu,\text{Fe}-\text{sim}}^{\text{det}}) - \ln(N_{\mu,p-\text{sim}}^{\text{det}})}$$



J. Albrecht et al., *The Muon Puzzle in cosmic-ray induced air showers and its connection to the Large Hadron Collider,* Astrophys.Space Sci. 367 (2022) 3, 27

Tuning of Free Parameters

Adjust free parameters to achieve a good description of data

Manual or brute-force tuning difficult due to high computing cost

Systematic event generator tuning workflow: Professor tuning system

arXiv:0907.2973

Tuning Systems

Professor: tune model parameters by numerical optimising the per-bin generator response (MINUIT) (*arXiv:0907.2973*)

Apprentice: improvement of the Professor tuning tool (*arXiv:2103.05748*)

Bayesian Analysis Monte Carlo Tuning

Bayesian Analysis MC tuning

Optimize free parameters of MC generator using experimental data and Bayesian Methods

PAPER • OPEN ACCESS

A Bayesian tune of the Herwig Monte Carlo event generator

Salvatore La Cagnina[1], Kevin Kröninger[1], Stefan Kluth[2] and Andrii Verbytskyi[2]

Published 26 October 2023 •

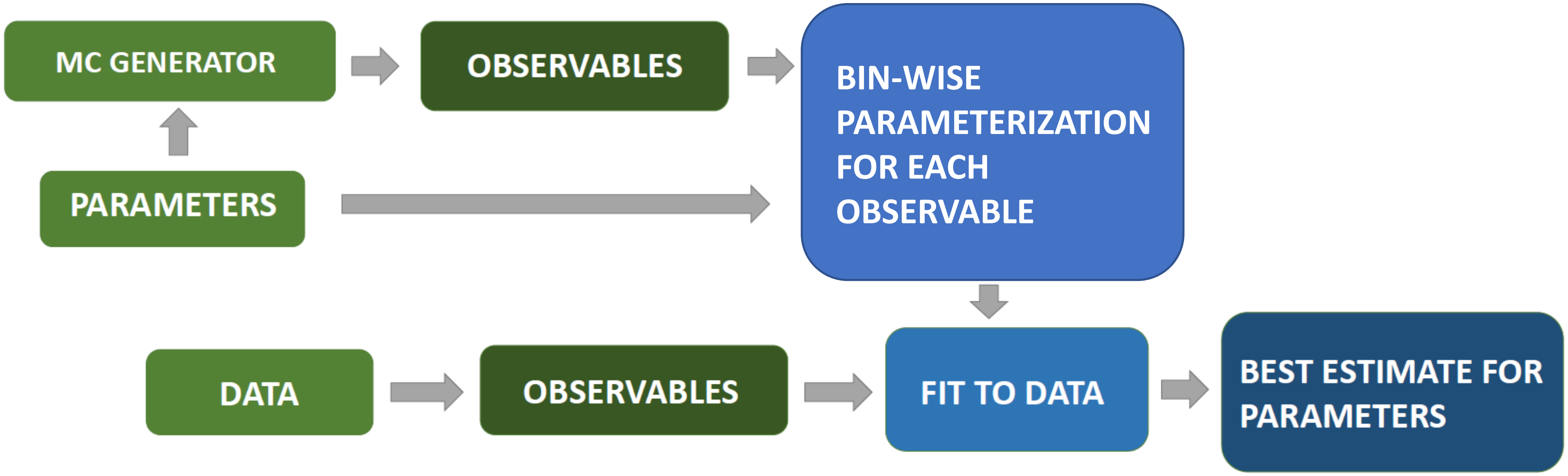
Journal of Instrumentation, Volume 18, October 2023

Citation Salvatore La Cagnina *et al* 2023 *JINST* **18** P10033

DOI 10.1088/1748-0221/18/10/P10033

Parameter based generator tuning

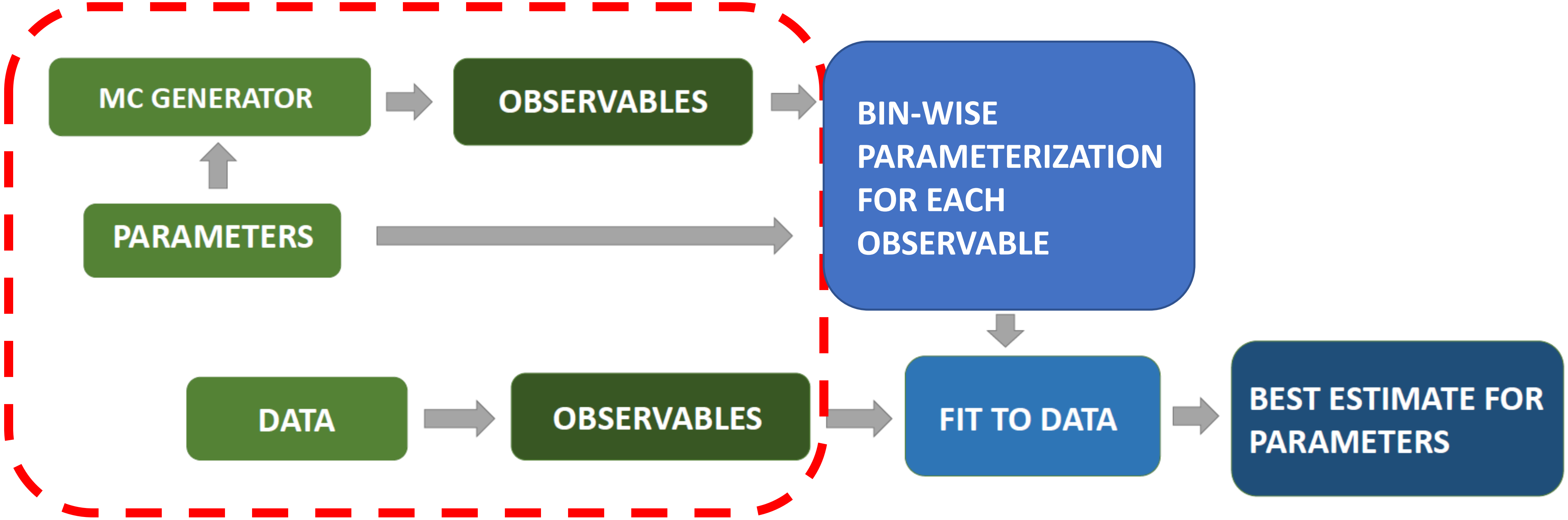
Optimize free parameters of MC generator using experimental data and Bayesian Methods



Parameter based generator tuning

Generate MC samples for different sets of parameter configurations

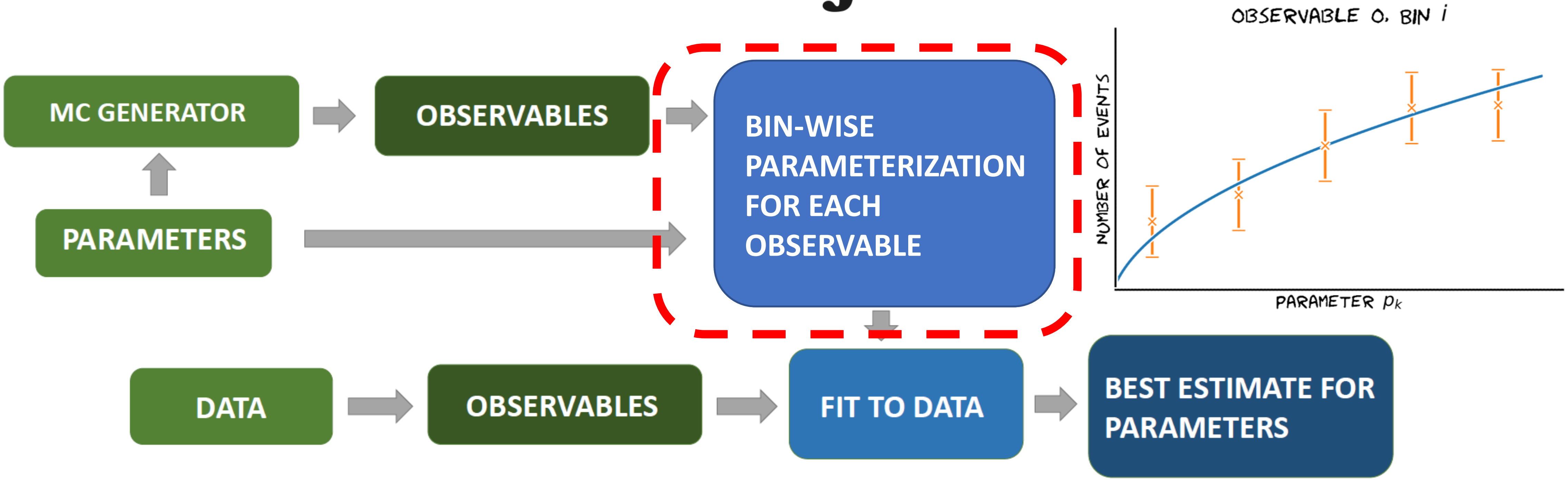
Reconstruct observables



Parameter based generator tuning

Create bin-wise parameterization of the observables as a function of the free parameters

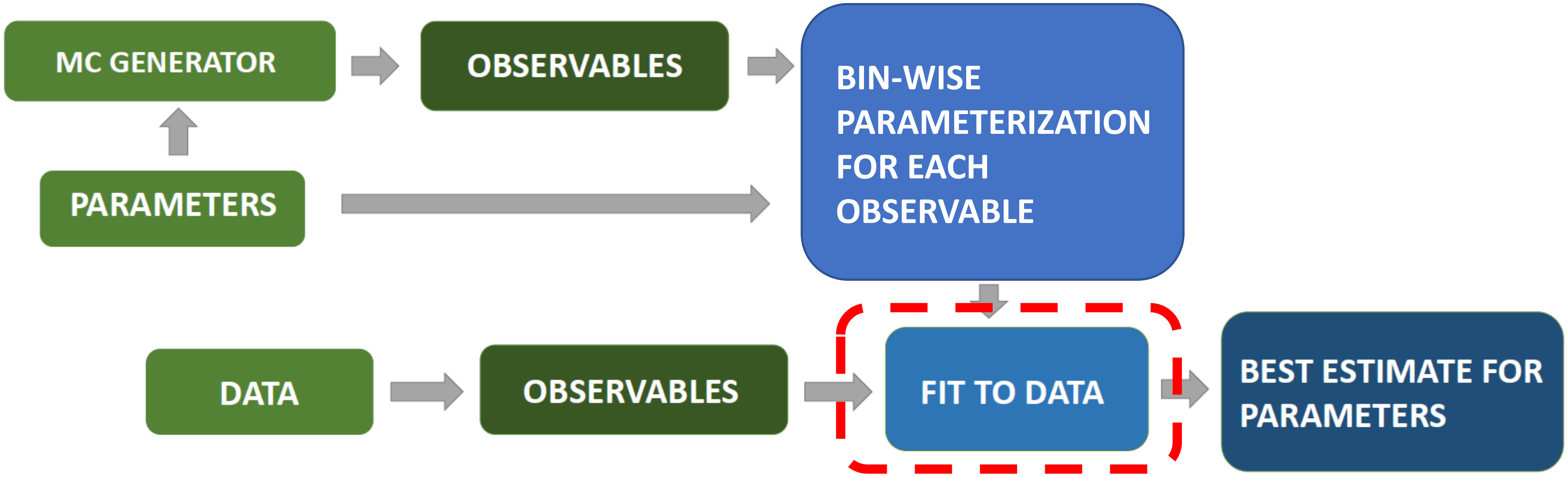
Written in julia



Parameter based generator tuning

Using the EFTfitter.jl tool a likelihood model is built from the parameterization and experimental data

N. Castro et al., *EFTfitter-A tool for interpreting measurements in the context of effective field theories,* Eur. Phys. J. C 76 (2016) 8, 432



EFTfitter.jl:

Parameters

Data

$$\ln L(\vec{D}|\vec{\lambda}) = -\frac{1}{2}[\vec{D} - \vec{f}(\vec{\lambda})]^T \cdot M^{-1} \cdot [\vec{D} - \vec{f}(\vec{\lambda})]$$

Likelihood

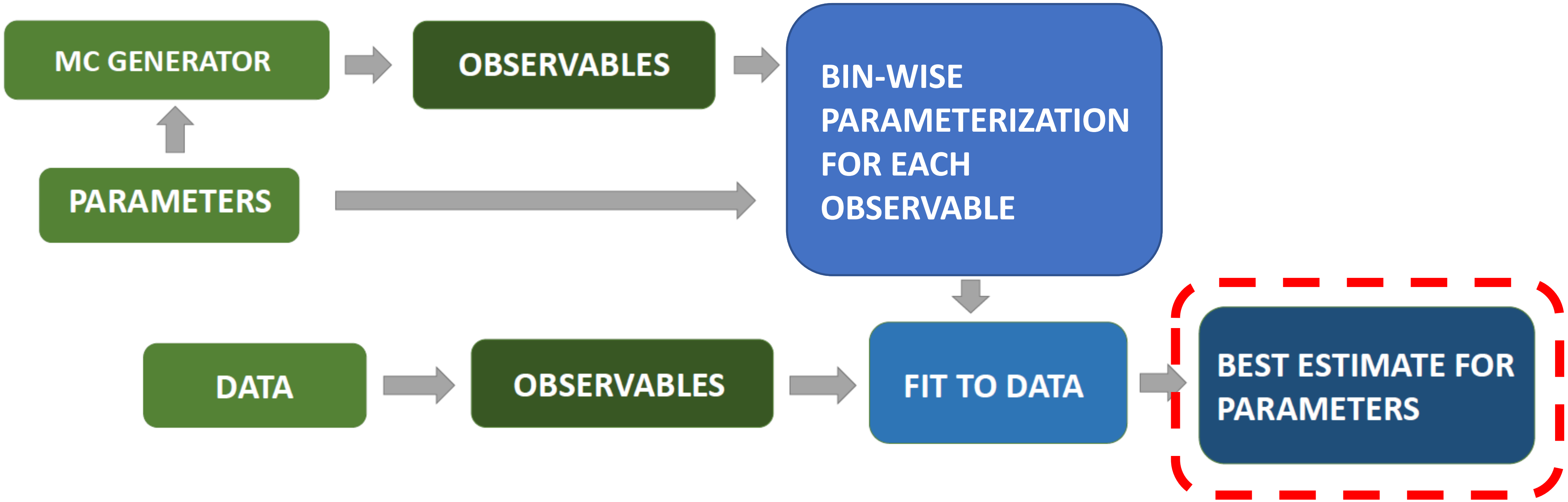
Parameterization

Covariance Matrix

Parameter based generator tuning

Using the BAT.jl framework the posterior space of the free parameters is sampled to achieve a tuned parameter setting

O. Schulz et al., *BAT.jl:A Julia-Based Tool for Bayesian Inference*, SNCS (2021)



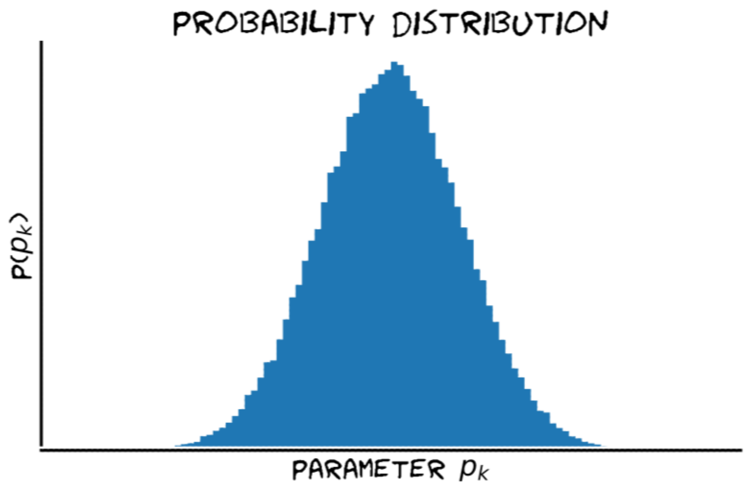
Sampling Posterior Space

Bayes' Theorem :

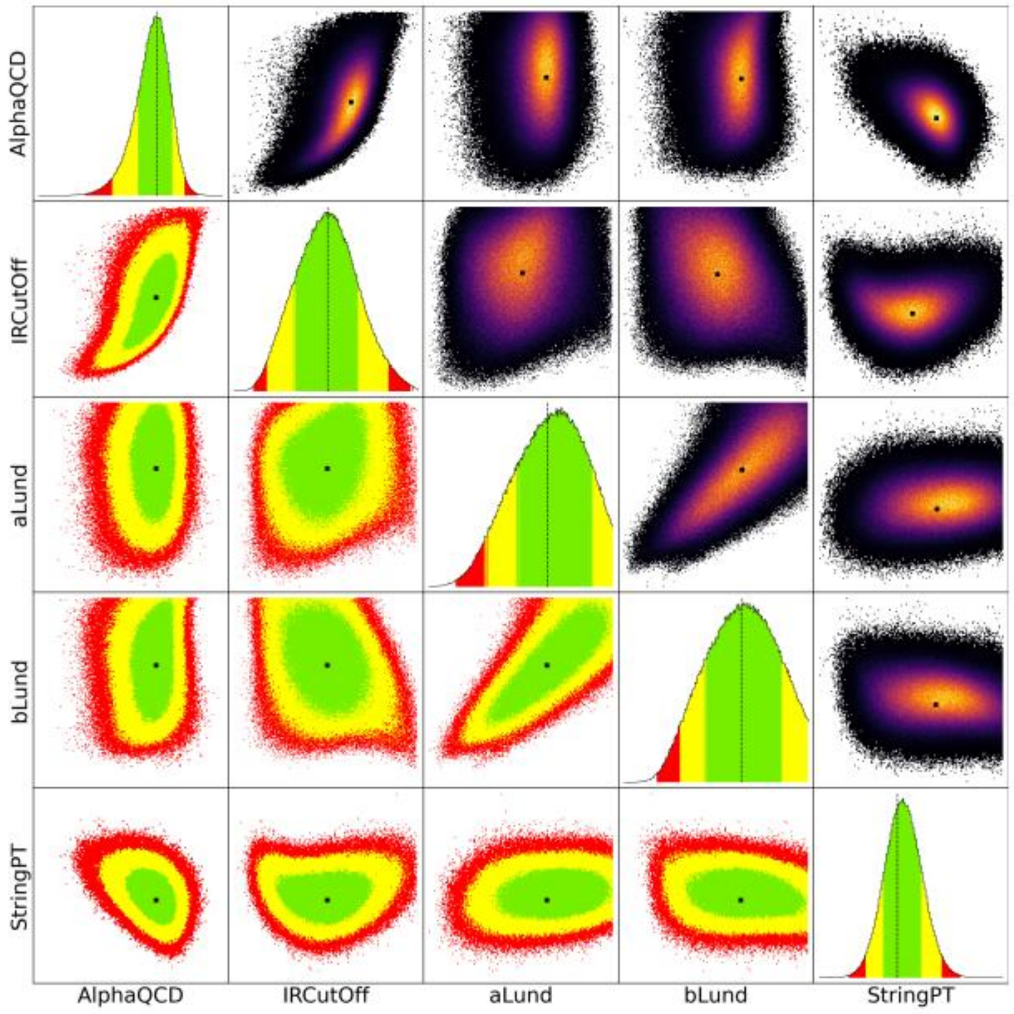
Posterior *Likelihood* *Prior*

$$p(\vec{\lambda}|\vec{D}) \propto L(\vec{D}|\vec{\lambda}) \cdot p(\vec{\lambda})$$

Markov Chain Monte Carlo
→ Metropolis-Hastings Algorithm (as default)



Works Well for Collider Data

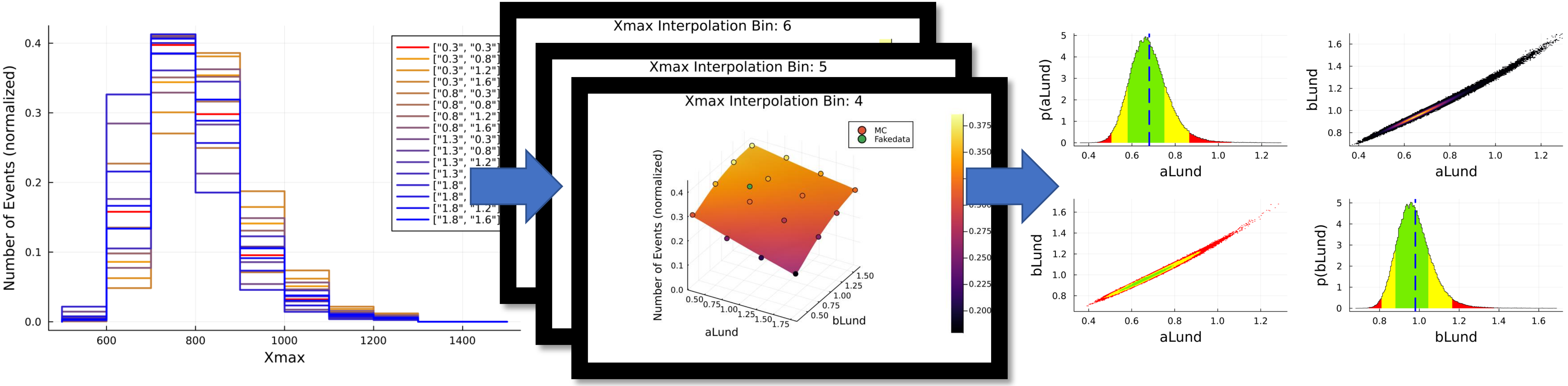


Toy Model For Pythia Tuning

Toy study based on PYTHIA example *main183*
→ Example includes simple study on air shower evolution in the atmosphere

T. Sjöstrand, M. Utheim, Hadron interactions for arbitrary energies and species, with applications to cosmic rays. Eur. Phys. J. C 82, 21 (2022)

Test Pythia tuning in toy model using only the Pythia 8 event generator



Air Shower Simulation and MC Generator

Choice of MC generator for tuning
→ PYTHIA: Collider background and convenient access to tuning

C. Bierlich et al., *A comprehensive guide to the physics and usage of PYTHIA 8.3.* (2022). arXiv:2203.11601

Use of air shower simulations

→CORSIKA8

→ Currently in development

→ PYTHIA as event generator

Simulation of CORSIKA8 samples

Pythia implementation in CORSIKA8 still in development

Simulating multiple sets of parameter configurations is cost heavy
→ Current bottleneck

Test by simulating one observable for multiple variations of one parameter
→ Remove electromagnetic component to reduce cost
→ Observable: Energy distribution of muons
→ Parameter: Strength of strong coupling constant α_s

CORSIKA8 Samples

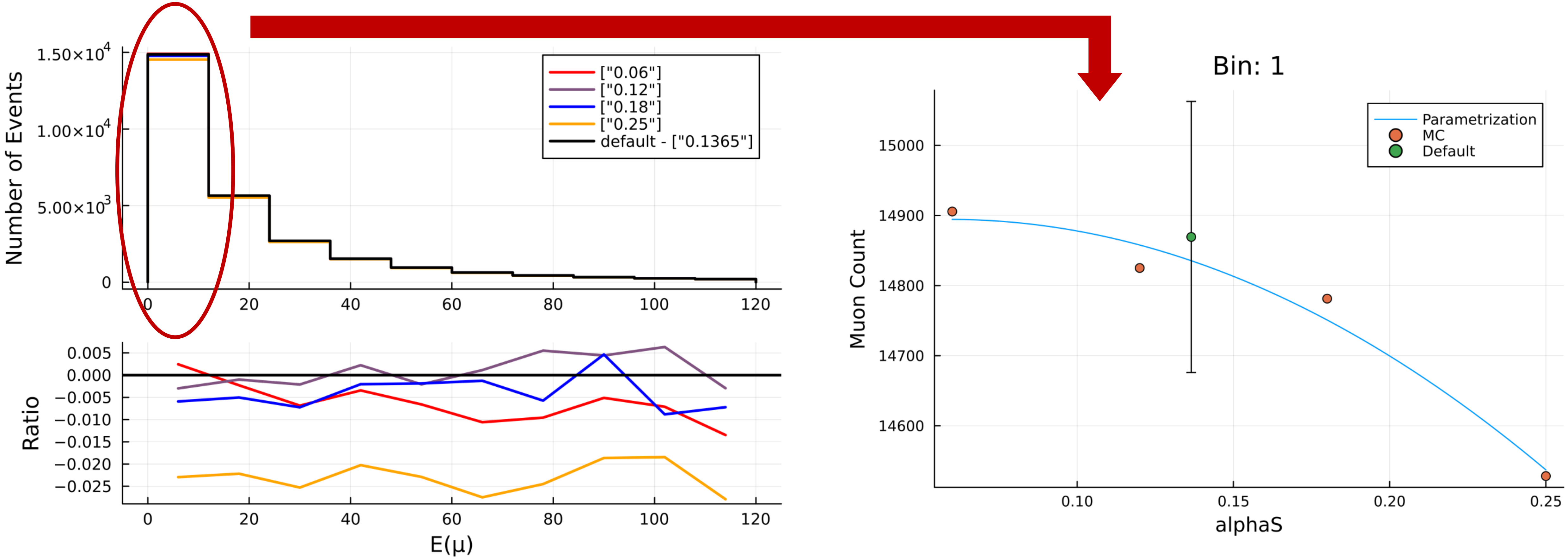
300 showers for each α_s configuration [0.06, 0.12, 0.18, 0.25]

→ Proton as primary particle with 10^7GeV

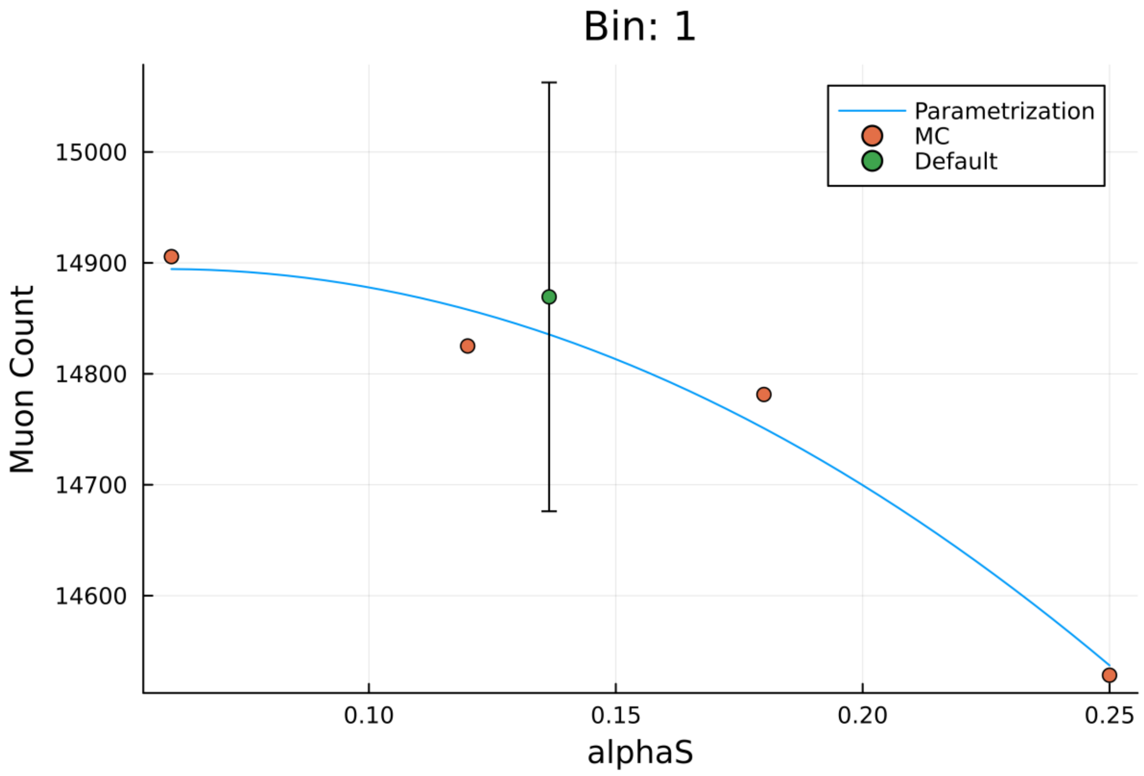
→ Lower hadronic cut: 0.5 GeV

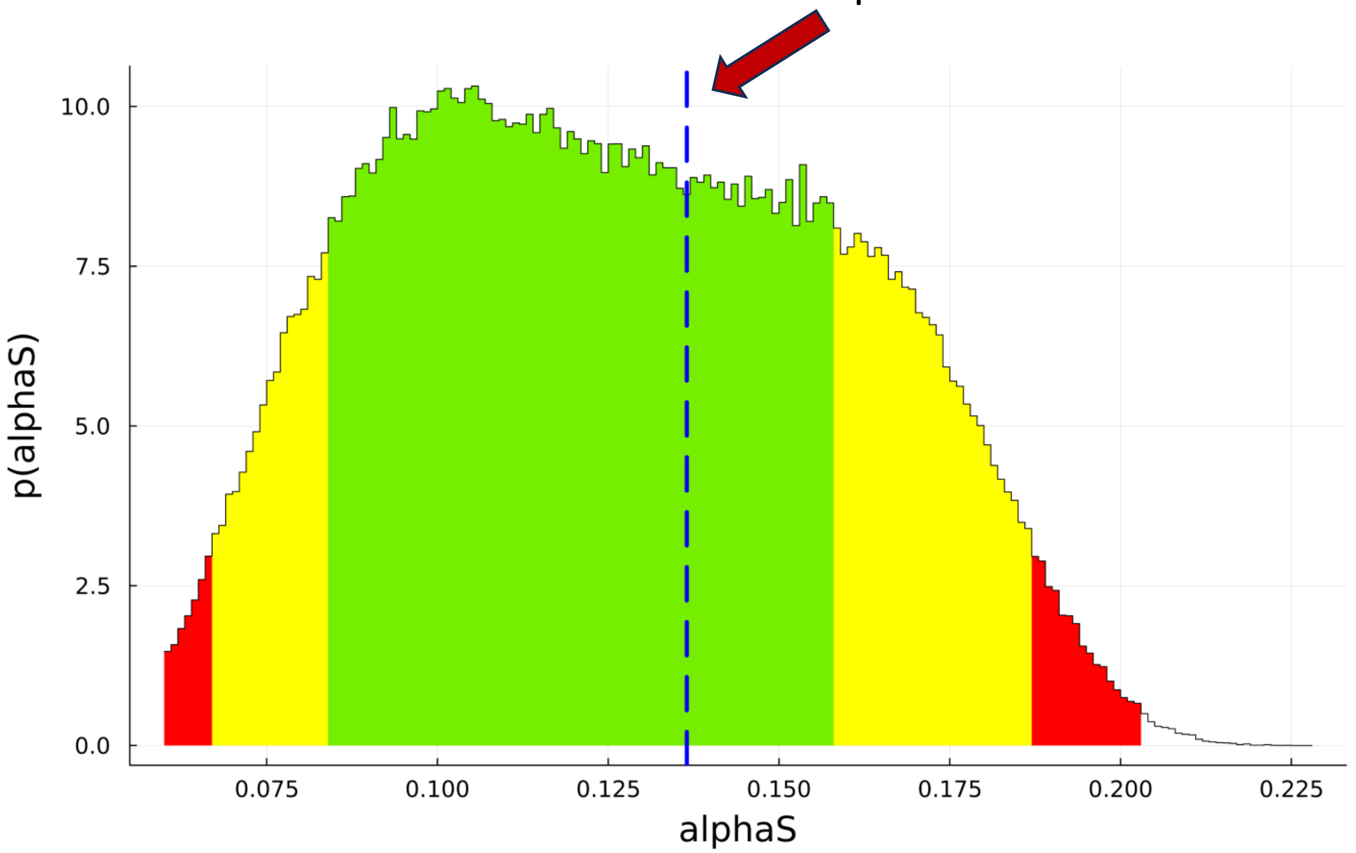
→ Lower muon cut: 0.3 GeV

CORSIKA8 Tuning – Need More Statistic



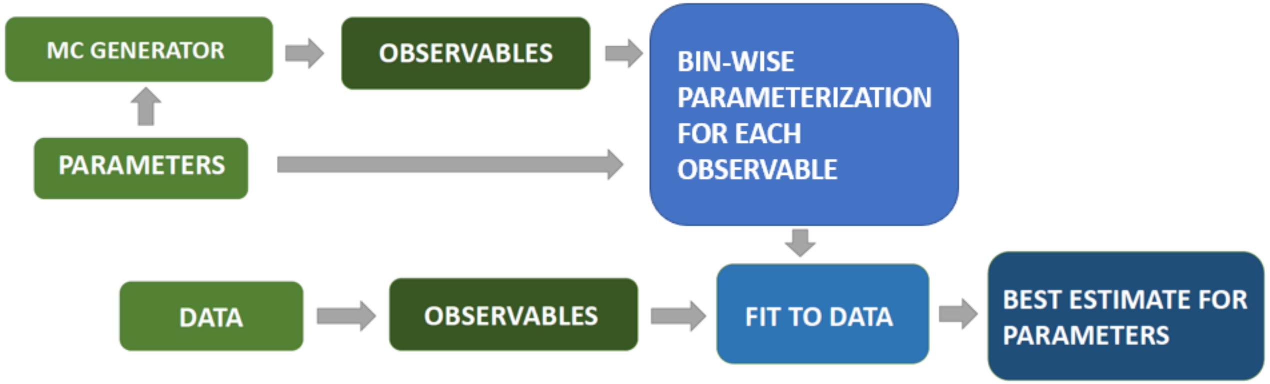
CORSIKA8 Tuning – Need More Statistic



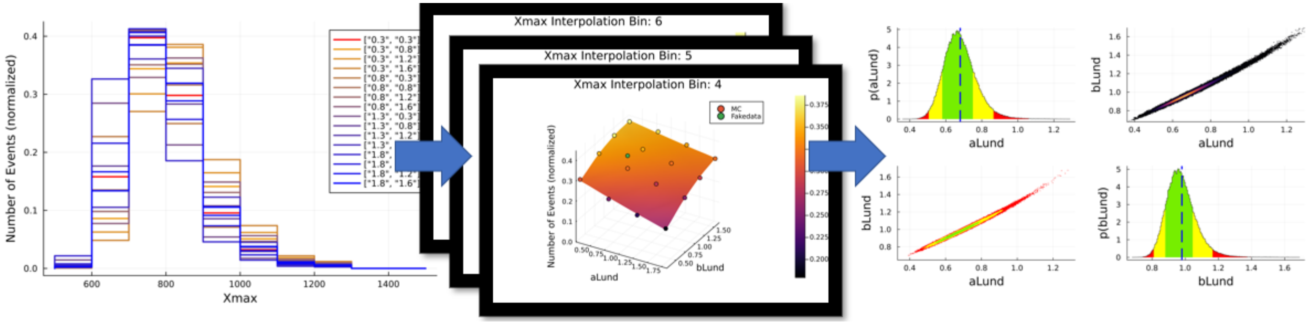


Conclude

Monte Carlo tuning using a Bayesian approach is possible



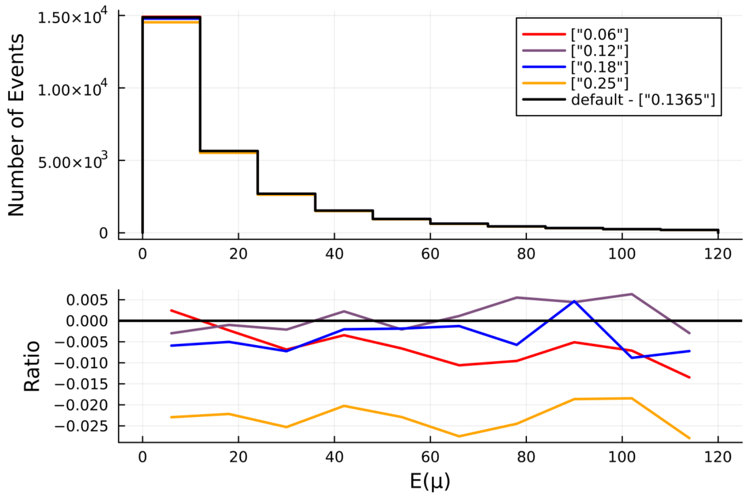
Tested on collider data for particle physics and in toy studies

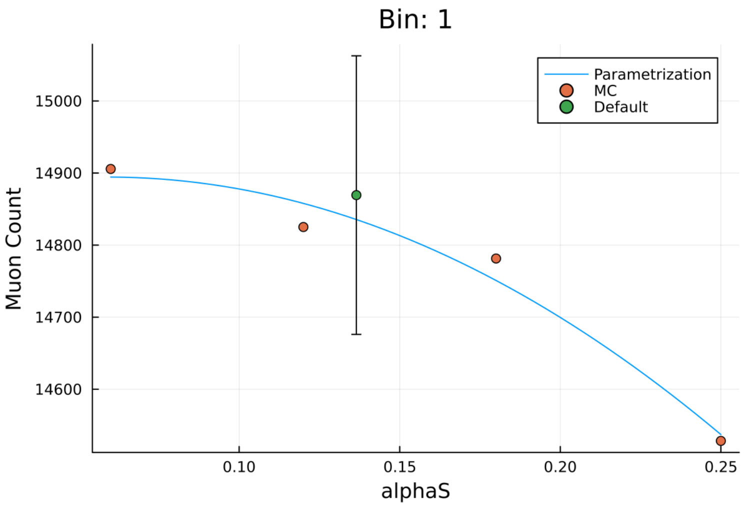


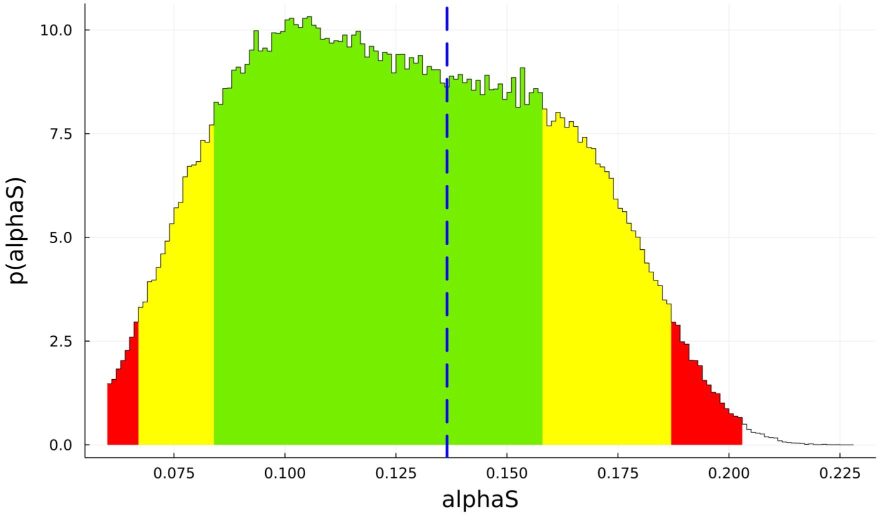
Outlook

Use of air shower data and CORSIKA8 for MC generator tuning

Combine collider and air shower data for simultaneous tune







Backup

Verify Tuning Algorithm

Simple toy study to test and demonstrate potential usefulness of tuning algorithm

Toy study based on PYTHIA example *main183*
→ Example includes simple study on air shower evolution in the atmosphere

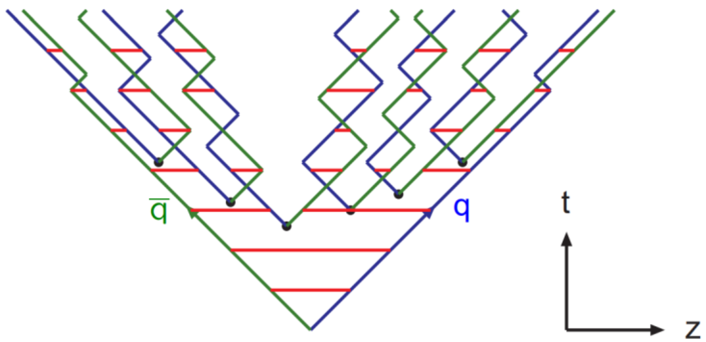
T. Sjöstrand, M. Utheim, Hadron interactions for arbitrary energies and species, with applications to cosmic rays. Eur. Phys. J. C 82, 21 (2022)

→ Provides quick and convenient generation of air shower observables for testing of tuning algorithm

Toy Study

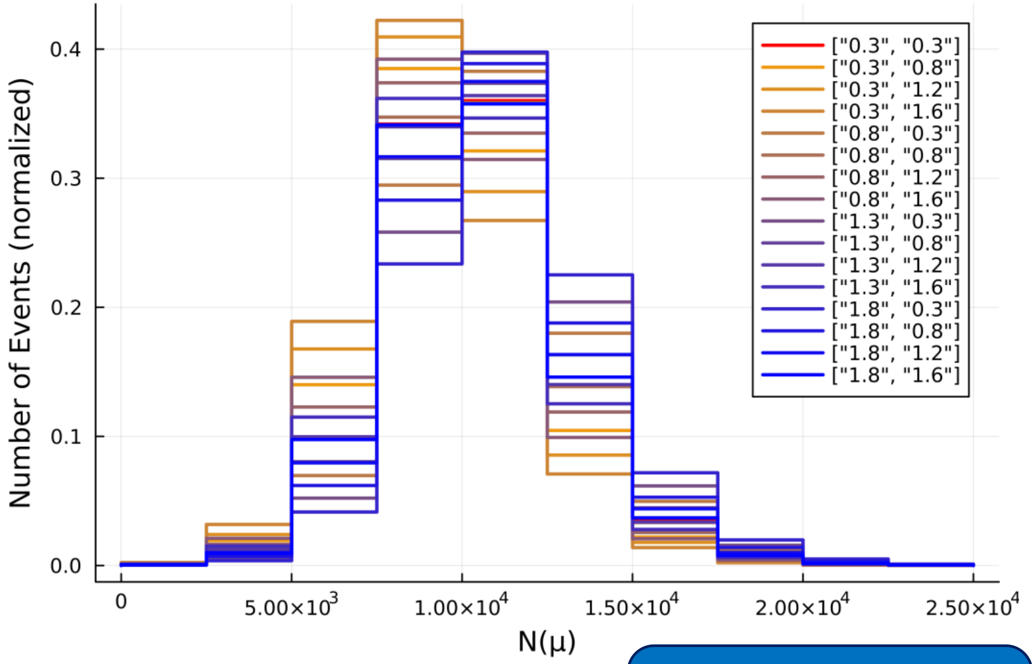
Study and test tuning method
→ Generate muon number and shower maximum observable
for different aLund and bLund parameter settings

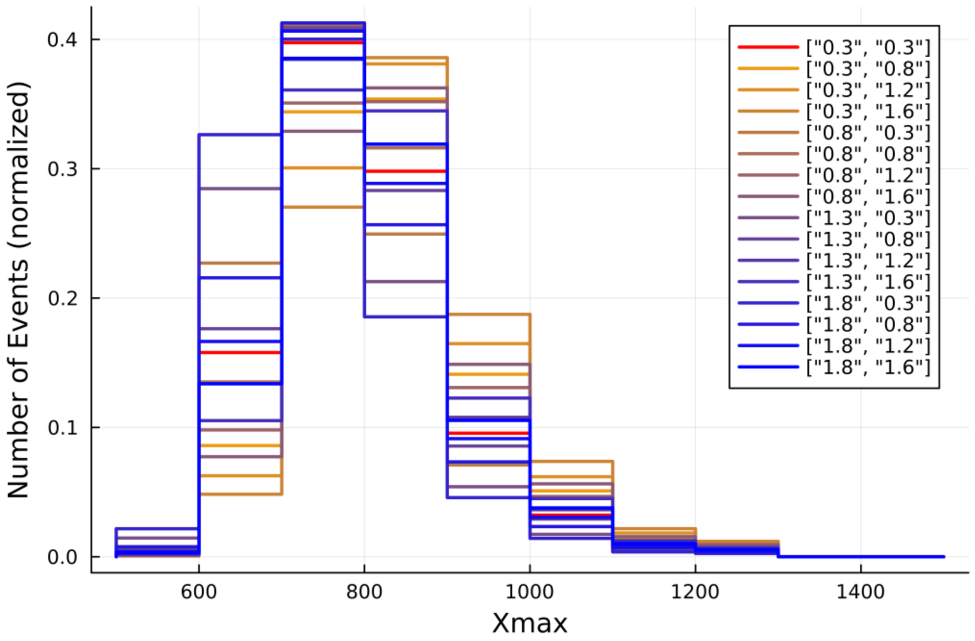
$$f(z) \propto z^{-1}(1-z)^{a} exp\left(\frac{-bm_{\perp}^{2}}{z}\right)$$

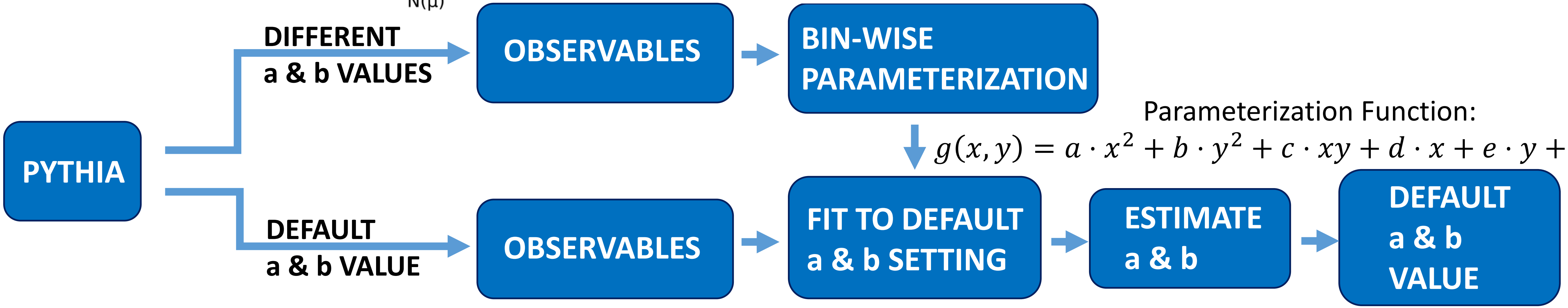


→ Use PYTHIAs default value for a- and bLund as *fake data* to tune against

Generate air shower simulations



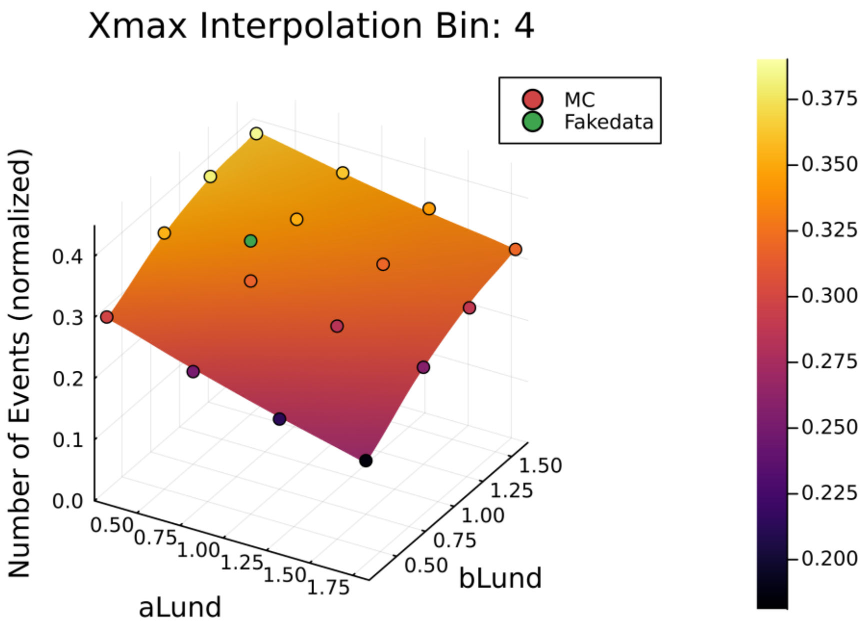


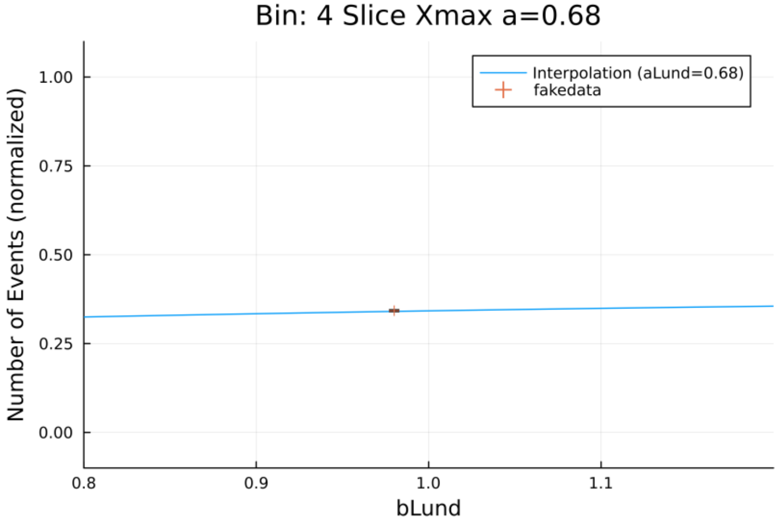


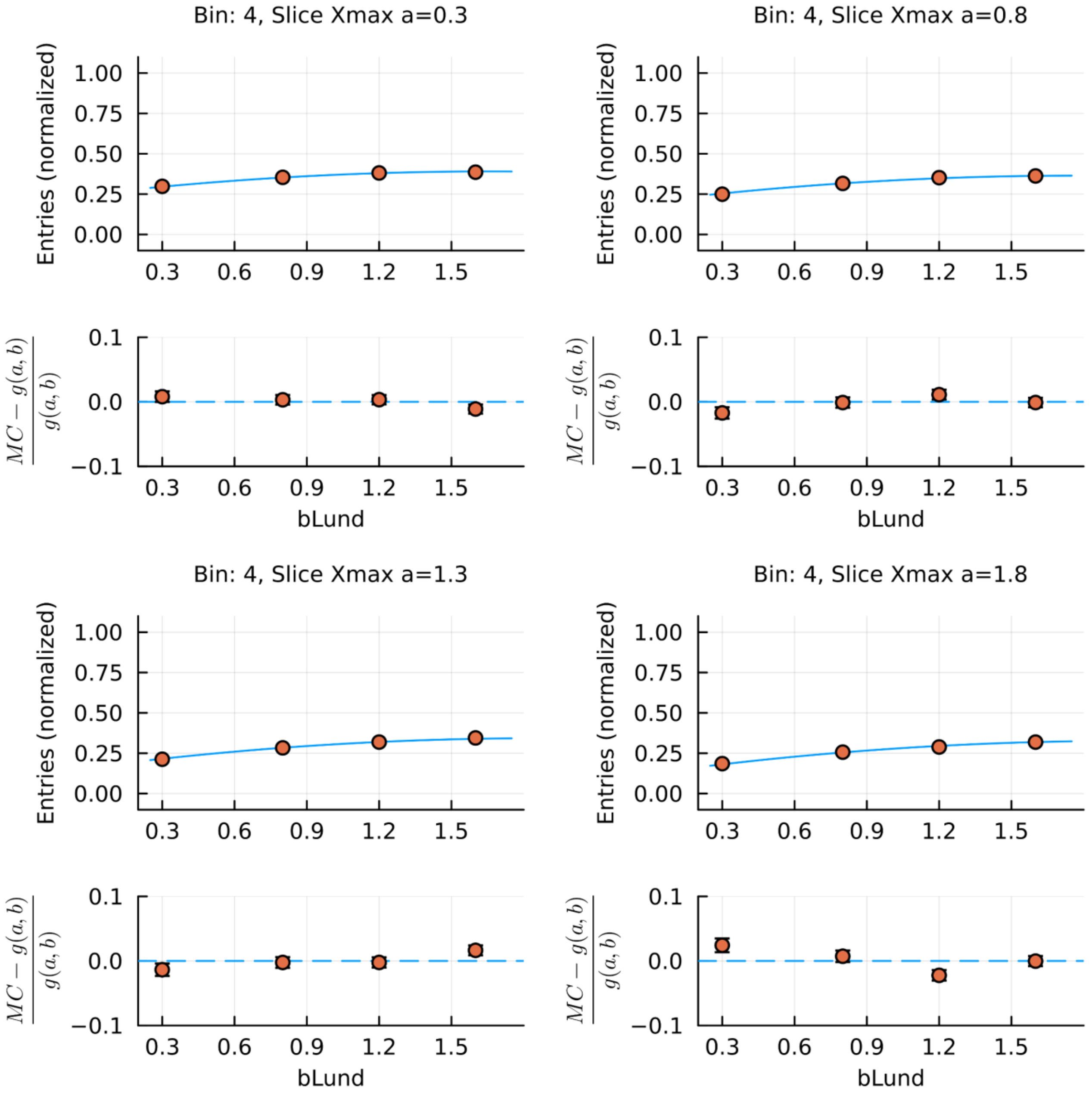
Parameterization Function:

$$g(x, y) = a \cdot x^2 + b \cdot y^2 + c \cdot xy + d \cdot x + e \cdot y + f$$

Parameterization







Sampling the posterior space